<u>REMARKS</u>

The present application was filed on October 31, 2000 with claims 1-27. Claims 1, 10 and 19 are the independent claims.

In the final Office Action, the Examiner: (i) rejects claims 1-27 based on 35 U.S.C. §112, second paragraph; (ii) rejects claims 1-27 under 35 U.S.C. §101 as being directed to non-statutory subject matter; (iii) rejects claims 1-8, 10-17 and 19-26 under 35 U.S.C. §102(b) as being anticipated by S. Chakrabarti et al., "Focused Crawling: A New Approach to Topic-Specific Web Resource Discovery," Computer Networks, 25 pages, 1999 (hereinafter "Chakrabarti"); and (iv) rejects claims 9, 18 and 27 as being unpatentable over Chakrabarti in view of S. Chakrabarti et al., "Distributed Hypertext Resource Discovery Through Examples," Proceedings of the 25 [th] VLDB Conference, Edinburgh, Scotland, pp. 375-386, 1999 (hereinafter "Ch2").

In this response, Applicants amend independent claims 1, 10 and 19, and traverse the §112, §101, §102(b) and §103(a) rejections for at least the following reasons.

Regarding the §112, second paragraph, rejections of claims 1-27, while Applicants believe that the previous amendment more than sufficiently satisfies §112, second paragraph, Applicants have nonetheless further defined the claimed invention in a sincere effort to expedite the case through to issuance.

Independent claim 1 now recites a computer-based method of performing document retrieval in accordance with an information network, the method comprising the steps of: initially retrieving one or more documents from the information network that satisfy a user-defined predicate, wherein the initial document retrieval operation is performed without assuming a specific model of a linkage structure such that the initial document retrieval operation retrieves the one or more documents without assuming that a relationship exists between a feature of a first one of the one or more documents and a feature of at least another one of the one or more documents that links to the first one; collecting statistical information about the one or more retrieved documents as the one or more retrieved documents are analyzed; and using the collected statistical information to automatically determine further document retrieval operations to be performed in accordance with the information network, wherein the statistical information using step further comprises learning a linkage structure from at least a portion of the collected statistical information with each successive document

retrieval operation such that the learned linkage structure is available for use in performing subsequent document retrieval operations requested by a user. Independent claims 10 and 19 recite similar limitations.

Again, while Applicants believe the previous claim language to be clear, concise, and fully supported by the specification, the independent claims have been amended to adopt the language that the Examiner uses to interpret the previous amendment, i.e., "without assuming a specific model of a linkage structure." Furthermore, in hopes of moving this case to issuance by removing any further concerns that may be raised by the Examiner, Applicants have also added further definition to the claim limitation by indicating that the claim language "without assuming a specific model of a linkage structure" at least means that "the initial document retrieval operation retrieves the one or more documents without assuming that a relationship exists between a feature of a first one of the one or more documents and a feature of at least another one of the one or more documents that links to the first one."

Support for the amendment may be found throughout the present specification. As illustratively explained in the present specification at page 4, line 22, through page 5, line 20:

> The present invention provides a more interesting and significantly more general alternative to conventional crawling techniques. As is evident from the teachings herein, no specific model for web linkage structure is assumed in intelligent crawling according to the invention. Rather, the crawler gradually learns the linkage structure statistically as it progresses. By linkage structure, we refer to the fact that there is a certain relationship between the content of a web page and the candidates that it links to. For example, a web page containing the word "Edmund Guide" is likely to link to web pages on automobile dealers. In general, linkage structure refers to the relationship between the various features of a web page such as content, tokens in Universal Resource Locators (URL), etc. Further, in general, it is preferred that the linkage structure be predicate-dependent. An intelligent crawler according to the invention learns about the linking structure during the crawl and find the most relevant pages. Initially, the crawler behavior is as random as a general crawler but it then gradually starts auto-focusing as it encounters documents which satisfy the predicate. A certain level of supervision in terms of documents which satisfy the predicate may be preferred since it would be very helpful in speeding up the process (especially for very specific predicates), but is not essential for the framework of the invention. This predicate may be a decision predicate or a quantitative predicate which assigns a certain level of priority to the search.
> The intelligent crawler of the invention may preferably be implemented as a graph search algorithm which works by treating web pages as nodes and links as edges. The

crawler keeps track of the nodes which it has already visited, and for each node, it decides the priority in which it visits based on its understanding of which nodes is likely to satisfy the predicate. Thus, at each point the crawler maintains candidate nodes which it is likely to crawl and keeps re-adjusting the priority of these nodes as its information about linkage structure increases .

Regarding the step of initially retrieving one or more documents from the information network that satisfy a user-defined predicate, as further support, the present specification, starting at page 8, line 22, explains that the input to the intelligent web crawling process of FIG. 2 includes a list of URLs to web pages from which the crawl starts, and a predicate which is used to focus and control the crawl. That is, there is no model for a linkage structure (e.g., as explained above) that is assumed in the initial step of the retrieval operation. A linkage structure is then gradually learned as the process progresses, as further illustrated in FIG. 2 and the subsequent figures.

In view of the above, Applicants respectfully request withdrawal of the §112, second paragraph, rejections of claims 1-27.

With regard to the §101 rejection of claim 1, Applicants respectfully point out that the Examiner's analysis of independent claim 1 under §101 makes some erroneous assumptions.

First, the Examiner states that the claim must recite a "practical application of a judicial exception." However, this assumes that claim 1 recites nothing more than an abstract idea or a law of nature. This is clearly not the case.

The claim recites tangible, concrete, useful results for a computer-based method of performing document retrieval in accordance with an information network. The method initially retrieves one or more documents from the information network that satisfy a user defined predicate. Such documents and user defined predicate can be considered as the object(s) being manipulated.

Based on statistical information collected about the one or more retrieved documents as the one or more retrieved documents are analyzed, the method automatically determines further document retrieval operations to be performed in accordance with the information network, wherein the statistical information using step further comprises learning a linkage structure from at least a portion of the collected statistical information with each successive document retrieval operation. Such further document retrieval operations and learned linkage structure can be considered concrete,

9

tangible, useful results of the method. The learned linkage structure is available for use in performing subsequent document retrieval operations made by a user.

To assist the Examiner in understanding how the further document retrieval operations and learned linkage structure can be considered concrete, tangible, useful results of the method, Applicants suggest that the present specification be reviewed to appreciate how, in one embodiment, such items are tangible, concrete and useful results of a web crawling methodology that improves upon conventional focused web crawling techniques.

Furthermore, the Examiner seems to contend that §101 requires a physical transformation or some sort of user presence. However, a review of the Computer-Related Invention guidelines in the MPEP will confirm that this is not the case. Nonetheless, the notion of physical manipulation is present in the claim since the retrieved documents are being manipulated in the retrieval and statistical information collection steps. The fact that the documents are in electronic form and being processed by a computer does not mean that there are no physical processes or transformations that occur. Also, user participation in the claimed methodology is evident in the use of the user defined predicate. Applicants have further amended the claim to indicate that the learned linkage structure is available for use in performing subsequent document retrieval operations made by a user.

Lastly, while recitation of a "practical application of a judicial exception" is not required in independent claim 1, a practical application is, in fact, recited. Document retrieval is most definitely a practical application, and is expressly recited throughout the claim.

Accordingly, withdrawal of the §101 rejection is respectfully requested.

With regard to the §102(b) rejection, Applicants initially note that MPEP §2131 specifies that a given claim is anticipated "only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference," citing Verdegaal Bros. v. Union Oil Co. of California, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). Moreover, MPEP §2131 indicates that the cited reference must show the "identical invention . . . in as complete detail as is contained in the . . . claim," citing Richardson v. Suzuki Motor Co., 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989). Applicants respectfully traverse the §102(b) rejection on the ground that the Chakrabarti reference fails to teach or suggest each and every limitation of claims 1-8, 10-17 and 19-26 as alleged.

Regarding the §102(b) rejection of claim 1, each and every one of the above-noted limitations of amended claim 1 fails to be anticipated by the teachings of Chakrabarti.

As Applicants noted in their previous response, the focused crawling approach of Chakrabarti employs an initial model where it is assumed that a specific linkage structure exists in which pages on a specific topic are likely to link to the same topic. That is, Chakrabarti assumes that there is a certain relationship between the content of a web page and the candidates that it links to. This is evident from the fact that Chakrabarti initiates crawling with a so-called "linkage sociology." For example, page 2, last paragraph of Chakrabarti refers to "discovering linkage sociology," in which examples of implementing the algorithm include inquiring: "is there a hyperlink between the web page of a speed trap (traffic radar) maker and an auto insurance company?...[a]part from other bicycling pages, what topics are prominent in the neighborhood of bicycling pages?... ([f]irst aid is one answer found by our system)." Also, page 8, second paragraph of Chakrabarti teaches, "the system starts by visiting all pages in $D(C^*)$...[i]n each step, the system can inspect its current set V of visited pages and then choose to visit an unvisited page from the crawl frontier, corresponding to a hyperlink on one or more visited pages...[i]nformally, the goal is to visit as many relevant pages and as few irrelevant pages as possible, i.e., to maximize average relevance."

That is, Chakrabarti discloses a method for focused crawling which includes making a decision to visit an unvisited page from the crawl frontier, corresponding to an initial link structure on one or more visited pages. Thus, Chakrabarti does not teach or suggest that an initial document retrieval operation is performed "without assuming a specific model of a linkage structure such that the initial document retrieval operation retrieves the one or more documents without assuming that a relationship exists between a feature of a first one of the one or more documents and a feature of at least another one of the one or more documents that links to the first one," as recited in the claimed invention, since the linkage sociology and crawl frontier assume that a relationship exists between a feature of a first document and a feature of at least another document that links to the first one.

The point is not that Chakrabarti may eventually visit "all pages in $D(C)^*$," but rather that Chakrabarti initially retrieves documents based on the assumption that a relationship exists between

a feature of a first document and a feature of at least another document that links to the first one, that is, pages on a specific topic are likely to link to pages on the same topic.

For at least the above reasons, Applicants respectfully assert that independent claims 1, 10 and 19 are patentable over Chakrabarti.

With regard to the § 103(a) rejection, Ch2 fails to supplement the deficiencies of Chakrabarti.

The remainder of the claims (namely, claims 2-9, 11-18 and 20-27) rejected over Chakrabarti depend, either directly or indirectly, from claims 1, 10 or 19, which are believed patentable for the reasons set forth above. Furthermore, the remaining claims define additional patentable subject matter in their own right.

Applicants point out that the final Office Action asserts that Applicants' have made certain acknowledgments and admissions in their previous response. However, Applicants respectfully contend that such assertions misrepresent or misunderstand Applicants' statements. It is believed that the above remarks will clarify any such issues.

In view of the above, Applicants respectfully request withdrawal of the § 112, § 101, § 102(b) and § 103(a) rejections of claims 1-27.

Respectfully submitted,

William E. Lewis
Attorney for Applicant(s)
Reg. No. 39,274
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560
(516) 759-2946

Date: February 22, 2007